ESTIMATION OF A MULTIVARIATE CONTINUOUS VARIABLE PANEL MODEL

TECHNICAL REPORT NO. 19

LAWRENCE S. MAYER AND KUN MAO CHEN

JUNE 1987

U. S. ARMY RESEARCH OFFICE
CONTRACT DAAG29-85-K-0239
THEODORE W. ANDERSON, PROJECT DIRECTOR

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

ESTIMATION OF A MULTIVARIATE CONTINUOUS VARIABLE PANEL MODEL*

Lawrence S. Mayer and Kun Mao Chen
Arizona State University

TECHNICAL REPORT NO. 19
June 1987

U. S. Army Research Office
Contract DAAG29-85-K-0239
Theodore W. Anderson, Project Director

Department of Statistics
Stanford University
Stanford, California

---

* The authors acknowledge discussions of the model analyzed with colleagues, particularly T. W. Anderson, and students. The research was completed, in part, while the first author was a Visiting Scholar, Department of Statistics, Stanford University.

1. Introduction

Panel studies are statistical studies in which two or more variables are observed for two or more subjects at two or more waves or points in time. In most panel studies the number of variables and the number of waves is small but the number of observations at any given wave is large. In this paper we focus on the model which has multiple responses, multiple predictors and an error structure that includes an individual specific random component. We consider the problem of computing the maximum likelihood estimates of the regression coefficients and covariance terms for this model and consider alternative estimators. We apply our results to a set of data on the relationship between the perceptions and attitudes of patients enrolled in a health maintenance organization.

The model we consider is closely related to panel models that have been studied widely in applied statistics and social methodology. Foremost it is a multivariate version of the "components of variance" model found in the statistics and econometrics literature [e.g., Balestra and Nerlove (1966), Chamberlain (1984), Hausman and Taylor (1981)]. Recently, excellent reviews of this work have become available [Dielman (1983), Hsiao (1985, 1986)].

Second, the model is both an extension and simplification of the most natural covariance structure formulation of the panel model, the model with independent errors over time [eg., Joreskog and Sorbom (1984), Joreskog and Goldberger (1975)]. It is an extension in that it allows errors to be correlated over time but a simplification in that it

assumes measurements are made without error and thus does not include a measurement model.

Finally, the model can be viewed as an alternative to the multivariate panel model with an autoregressive error structure over waves [e.g., Mayer (1986), Mayer and Carroll (1986)]. The autoregressive-error model assumes that individuals have a common error structure over time and that for a single individual the correlation between the errors at two waves is a geometrically decreasing function of the time separating the waves. The individual effect model we consider assumes that the temporal structure of the errors differs across individuals but that for a single individual the correlation between the errors at two waves does not depend on the time between the waves.

The results we offer are heavily dependent on the results of T. W. Anderson (1969, 1970, 1973, 1984) on estimating covariance matrices with a known structure and his results on vector-valued autoregressive models [Anderson (1978)], on the comprehensive study of T. Szatrowski (1980) of necessary and sufficient conditions for a linear model to yield closed form maximum likelihood estimators, and, finally, on the work of Magnus (1982) on maximum likelihood estimation in a multivariate panel model somewhat similar to the model we examine. Our interest in using this model to analyze social science data was inspired, in part, by the work of Anderson and Hsiao (1981, 1982) on maximum likelihood estimation in a univariate version of this model which includes lagged endogenous predictors, although, in this paper we do not tackle the problem of lagged endogenous predictors.

We close the introduction by noting that since the individual effect panel model can be cast as a covariance structure model, values of the maximum likelihood estimates are readily obtained by applying a standard covariance structure estimation program (provided it is capable of handling somewhat complicated constraints on the covariance matrix). Thus the model can be estimated using procedures that have become fairly standard in social science research; in fact, we use such a program (LISREL) to obtain estimates. Our role as statisticans serving social scientists is to study the behavior of the statistical procedures, including estimators, being used in empirical social research. In the spirit of this service we offer our results.

In the next section we formulate the model and dispense with our preliminaries; in the third, we present our results on estimation for our model, and in the fourth we present a numerical example of our calculations. The last section contains brief comments on our results and current research directions.

II. Preliminaries

Let $\underset{\sim}{z}_{it}$ be a p dimensional response vector for n independent subjects at T waves. The structure of the response is

$$\underset{\sim}{z}_{it} = \underset{\sim}{B}\,\underset{\sim}{w}_{it} + \underset{\sim}{e}_{it} \quad i = 1,\ldots,n \; ; \; t = 1,\ldots,T \tag{2.1}$$

where $\underset{\sim}{w}_{it}$ is a k dimensional exogenous vector, $\underset{\sim}{B}$ is a p x k matrix of regression coefficients, and $\underset{\sim}{e}_{it}$ is a p dimensional unobserved error vector, uncorrelated with $\underset{\sim}{w}_{it}$, and with structure

$$\underset{\sim}{e}_{it} = \underset{\sim}{\alpha}_i + \underset{\sim}{u}_{it} \quad i = 1,\ldots,n \; ; \; t = 1,\ldots,T \tag{2.2}$$

where α_i and u_{it} are independent normal (Gaussian) errors, with common mean 0 and common nonsingular covariance matrices Λ and Σ, respectively. Provided $i \neq j$ or $t \neq s$ α_i and α_j are independent as are u_{it} and u_{js}.

We re-express, in vector form, the model by letting $z_i = (z_{i1}', \ldots, z_{iT}')'$, $w_i = (w_{i1}', \ldots, w_{iT}')'$ and $e_i = (e_{i1}', \ldots, e_{iT}')'$ then the equation in (2.1) becomes

$$z_i = (I \otimes B)\, w_i + e_i \tag{2.3}$$

If we let Ω be the covariance matrix of e_i then

$$\Omega = I \otimes \Sigma + 11' \otimes \Lambda \tag{2.4}$$

where $\otimes$ indicates the Kronecker product. The diagonal blocks of Ω are identically $\Sigma + \Lambda$ and the off-diagonal blocks are identically Λ.

In anticipation of developing results similar to those of Anderson (1973) we reexpress the model as

$$z_i = X_i\, b + e_i \tag{2.5}$$

where X_i is a known matrix and b is a unknown vector.

To effect this reexpression we write B and Ω in the form

$$B = \sum_{n=1}^{r} \beta_h H_h \qquad\qquad \Omega = \sum_{g=0}^{m} \sigma_g G_g$$

where the H_h's and G_g's are known matrices, the β_h's and σ_g's are unknown scalars, r is the number of distinct elements in B and m + 1

is the number of distinct elements in $\underset{\sim}{\Omega}$. Continuing, let $\underset{\sim}{x}_{hi} =$ vec $(\underset{\sim}{H}_h \underset{\sim}{W}_i)$ where $\underset{\sim}{W}_i = (\underset{\sim}{w}_{i1}, \ldots, \underset{\sim}{w}_{iT})$; let $\underset{\sim}{X}_i = (\underset{\sim}{x}_{1i}, \ldots, \underset{\sim}{x}_{ri})$; and let $\underset{\sim}{b} = (\beta_1, \ldots, \beta_r)'$. Using $\underset{\sim}{X}_i$ and $\underset{\sim}{b}$ as just defined, the expression in (2.5) is obtained by applying the following result of Neudecker (1969):

<u>Lemma 1</u>: For any conformal matrices $\underset{\sim}{X}$, $\underset{\sim}{Y}$ and $\underset{\sim}{Z}$

$$\text{vec}\ (\underset{\sim}{X}\underset{\sim}{Y}\underset{\sim}{Z}) = (\underset{\sim}{Z}' \otimes \underset{\sim}{X})\ \text{vec}\ (\underset{\sim}{Y})$$

where if $\underset{\sim}{A}$ is a n x p matrix with ith row $\underset{\sim}{a}_i$ then vec $(\underset{\sim}{A})$ is the vector $(\underset{\sim}{a}_1', \ldots, \underset{\sim}{a}_n')'$.

III. Maximum Likelihood Estimation

Our results begin by applying the scoring method of Anderson (1973) to obtain implicit equations for the maximum likelihood estimates. We offer

<u>Proposition 1</u>: The maximum likelihood estimates of $\underset{\sim}{B}$, $\underset{\sim}{\Sigma}$, and $\underset{\sim}{\Lambda}$ are the solutions of the equations

$$\text{tr}\left(\sum_{h=0}^{m} \sigma_h \underset{\sim}{G}_h\right)^{-1} \underset{\sim}{G}_g = \text{tr}\left(\sum_{h=0}^{m} \sigma_h \underset{\sim}{G}_h\right)^{-1} \underset{\sim}{G}_g \left(\sum_{h=0}^{m} \sigma_h \underset{\sim}{G}_h\right)^{-1} \underset{\sim}{C} \qquad (3.1)$$

where

$$\underset{\sim}{C} = \frac{1}{n} \sum_{i=1}^{n} [\underset{\sim}{z}_i - \text{vec}\ (\underset{\sim}{B}\underset{\sim}{W}_i)][\underset{\sim}{z}_i - \text{vec}\ (\underset{\sim}{B}\underset{\sim}{W}_i)]'$$

and

$$\sum_{i=1}^{n} \sum_{j=1}^{r} \underset{\sim}{x}_{hi}' \underset{\sim}{\Omega}^{-1} \underset{\sim}{x}_{ji} \beta_j = \sum_{i=1}^{n} \underset{\sim}{x}_{hi}' \underset{\sim}{\Omega}^{-1} \underset{\sim}{z}_i \tag{3.2}$$

The proof of Proposition 1 begins with the likelihood function of all Tnk observations which can be espressed as

$$L(\underset{\sim}{B}, \underset{\sim}{\Omega}) = (2\pi)^{-nkT/2} |\underset{\sim}{\Omega}|^{-n/2} \exp\{-\frac{1}{2} \sum_{i=1}^{n} [\underset{\sim}{z}_i - (\underset{\sim}{I} \otimes \underset{\sim}{B}) \underset{\sim}{w}_i]' \underset{\sim}{\Omega}^{-1}$$

$$[\underset{\sim}{z}_i - (\underset{\sim}{I} \otimes \underset{\sim}{B}) \underset{\sim}{w}_i]\}$$

$$\text{Let } \ell = (2/n) \log L = c_1 - \log |\underset{\sim}{\Omega}| - \mathrm{tr}\, \underset{\sim}{\Omega}^{-1} \underset{\sim}{C} \tag{3.3}$$

where c_1 is the constant $-kT \log(2\pi)$ and

$$\underset{\sim}{C} = n^{-1} \sum_{i=1}^{n} [\underset{\sim}{z}_i - (\underset{\sim}{I} \otimes \underset{\sim}{B}) \underset{\sim}{w}_i][\underset{\sim}{z}_i - (\underset{\sim}{I} \otimes \underset{\sim}{B}) \underset{\sim}{w}_i]' \tag{3.4}$$

Define the matrix $\underset{\sim}{Y}_i = (\underset{\sim}{z}_{i1}, \ldots, \underset{\sim}{z}_{iT})$. Since $(\underset{\sim}{I} \otimes \underset{\sim}{B}) \underset{\sim}{w}_i = \mathrm{vec}\,(\underset{\sim}{B}\underset{\sim}{W}_i)$ and $\underset{\sim}{z}_i = \mathrm{vec}\,(\underset{\sim}{Y}_i)$, $\underset{\sim}{C}$ can be expressed as

$$\underset{\sim}{C} = n^{-1} \sum_{i=1}^{n} [\underset{\sim}{z}_i - \underset{\sim}{\mu}_i][\underset{\sim}{z}_i - \underset{\sim}{\mu}_i]' \tag{3.5}$$

where $\mu_i = \mathrm{vec}\,(\underset{\sim}{B}\underset{\sim}{W}_i)$.

Applying to ℓ the following formulae of Anderson (1973)

$$(\partial \underset{\sim}{\Omega} / \partial \sigma_g) = \underset{\sim}{G}_g$$

$$(\partial \underset{\sim}{\Omega}^{-1} / \partial \sigma_g) = -\underset{\sim}{\Omega}^{-1} \underset{\sim}{G}_g \underset{\sim}{\Omega}^{-1}$$

$$\frac{\partial \log |\underset{\sim}{\Omega}|}{\partial \sigma_g} = \operatorname{tr} \underset{\sim}{\Omega}^{-1} \underset{\sim}{G}_g$$

yields

$$\frac{\partial \ell}{\partial \sigma_g} = -\operatorname{tr} \underset{\sim}{\Omega}^{-1} \underset{\sim}{G}_g + \operatorname{tr}(\underset{\sim}{\Omega}^{-1} \underset{\sim}{G}_g \underset{\sim}{\Omega}^{-1}) \underset{\sim}{C} \tag{3.6}$$

substituting vec $(\sum_{h=1}^{n} \beta_h \underset{\sim}{H}_h \underset{\sim}{W}_i)$ for $\underset{\sim}{\mu}_i$ in equation (3.5) and differentiating equation (3.3) with respect to β_h yields

$$\frac{\partial \ell}{\partial \beta_h} = \frac{-2}{n} \sum_{i=1}^{n} [-\text{vec } (\underset{\sim}{H}_h \underset{\sim}{W}_i)]' \underset{\sim}{\Omega}^{-1} [\text{vec}(\underset{\sim}{Y}_i) - \text{vec}(\underset{\sim}{B}\underset{\sim}{W}_i)] \tag{3.7}$$

Setting the derivatives equal to zero yields the estimating equations

$$\operatorname{tr} \underset{\sim}{\Omega}^{-1} \underset{\sim}{G}_g = \operatorname{tr} \underset{\sim}{\Omega}^{-1} \underset{\sim}{G}_g \underset{\sim}{\Omega}^{-1} \underset{\sim}{C} \tag{3.8}$$

$$\sum_{i=1}^{n} [\text{vec } (\underset{\sim}{H}_h \underset{\sim}{W}_i)]' \underset{\sim}{\Omega}^{-1} [\text{vec } (\underset{\sim}{B}\underset{\sim}{W}_i)] =$$

$$\sum_{i=1}^{n} [\text{vec } (\underset{\sim}{H}_h \underset{\sim}{W}_i)]' \underset{\sim}{\Omega}^{-1} [\text{vec } (\underset{\sim}{Y}_i)] \tag{3.9}$$

since $\underset{\sim}{\Omega} = \sum_{h=0}^{m} \sigma_h \underset{\sim}{G}_h$ equation (3.8) can be expressed as equation (3.1).

Since vec $(\underset{\sim}{B}\underset{\sim}{W}_i) = \sum_{j=1}^{r} \text{vec}(\underset{\sim}{H}_i \underset{\sim}{W}_i) \beta_j$ equation (3.9) can be expressed as

$$\sum_{i=1}^{n} \sum_{j=1}^{r} [\text{vec}(\underset{\sim}{H}_h \underset{\sim}{W}_i)]' \underset{\sim}{\Omega}^{-1} [\text{vec}(\underset{\sim}{H}_j \underset{\sim}{W}_i) \beta_i] =$$

$$\sum_{i=1}^{n} [\mathrm{vec}(\underset{\sim}{H}_h \underset{\sim}{W}_i)]' \; \underset{\sim}{\Omega}^{-1} \; [\mathrm{vec}(\underset{\sim}{Y}_i)] \tag{3.10}$$

Since $\underset{\sim}{x}_{hi} = \mathrm{vec}\ (\underset{\sim}{H}_h \underset{\sim}{W}_i)$, equation (3.10) becomes

$$\sum_{i=1}^{n} \sum_{j=1}^{r} \underset{\sim}{x}_{hi}' \; \underset{\sim}{\Omega}^{-1} \; \underset{\sim}{x}_{hi} \; \beta_j = \sum_{i=1}^{n} \underset{\sim}{x}_{hi}' \; \underset{\sim}{\Omega}^{-1} \; \underset{\sim}{z}_i$$

which is equation (3.2) and the result is complete.

Following the suggestion of Anderson (1973) the estimating equations [(3.1) and (3.2)] can be used to iterate initial estimates of $\underset{\sim}{B}$, $\underset{\sim}{\Lambda}$, and $\underset{\sim}{\Sigma}$. Begin by setting the covariance matrix $\underset{\sim}{\Omega}$ to the identity matrix and solving equation (3.2) for initial estimates $\beta_1^{(0)}, \ldots, \beta_r^{(0)}$ of $\beta_1, \ldots, \beta_r$. These estimates are used to give an initial estimate $\underset{\sim}{B}^{(0)} = \sum_{j=1}^{r} \beta_j^{(0)} \underset{\sim}{H}_j$ of $\underset{\sim}{B}$. A pT x pT matrix $\underset{\sim}{C}^{(0)}$ is then computed by substituting $\underset{\sim}{B}^{(0)}$ for $\underset{\sim}{B}$ in equation (3.4). The matrix $\underset{\sim}{C}^{(0)}$ is substituted for $\underset{\sim}{C}$ in the expression

$$\sum_{f=0}^{m} \mathrm{tr}\ \underset{\sim}{\Theta}\ \underset{\sim}{G}_g\ \underset{\sim}{\Theta}\ \underset{\sim}{G}_f\ \hat{\sigma}_f = \mathrm{tr}\ \underset{\sim}{\Theta}\ \underset{\sim}{G}_g\ \underset{\sim}{\Theta}\ \underset{\sim}{C} \tag{3.11}$$

where $\underset{\sim}{\Theta}$ is a fixed matrix. The equation in (3.11) is solved for initial estimates $\hat{\sigma}_f^{(0)}$ which are used to give an initial estimate of the covariance matrix $\underset{\sim}{\Omega}$,

$$\underset{\sim}{\Omega}^{(0)} = \sum_{f=0}^{m} \hat{\sigma}_f^{(0)} \underset{\sim}{G}_f$$

The procedure continues by substituting $\underset{\sim}{\Omega}^{(0)}$ for $\underset{\sim}{\Omega}$ in (3.2) and solving

for new estimates $\beta_1^{(1)},\ldots,\beta_r^{(1)}$ and thus for a new estimate $\underset{\sim}{B}^{(1)}$ of $\underset{\sim}{B}$. This new estimate is used in the equation in (3.5) to yield $\underset{\sim}{C}^{(1)}$ which is used for $\underset{\sim}{C}$ in equation (3.1), and $\hat{\sigma}_F^{(1)}$ is computed. Then $\underset{\sim}{\Omega}^{(1)} = \sum_{f=0}^{m} \hat{\sigma}_f^{(1)} \underset{\sim}{G}_f$ is obtained. Continuing until convergence yields the maximum likelihhod estimates. The matrix $\underset{\sim}{\Theta}$ in (3.11) may be chosen to simplify calculations. We use $\underset{\sim}{\Theta} = \underset{\sim}{I}$.

Standard multivariate statistical theory yields that the maximum likelihood estimator $\hat{\underset{\sim}{B}}$ is asymptotically efficient and asymptotically normal with covariance matrix that has

$$[1/2 \ \mathrm{tr}\ \underset{\sim}{\Omega}^{-1}\underset{\sim}{G}_h\underset{\sim}{\Omega}^{-1}\underset{\sim}{G}_g]^{-1} \tag{3.12}$$

as its (h,g)th entry. If only one iteration is used we call the resultant estimator the Anderson "one-step estimator." The usefulness of the one-step estimators arises from a result of Anderson (1973) which guarantees that they enjoy the asymptotic properties given above for the maximum likelihood estimator.

In practice, these equations [(3.1) and (3.2)] are most useful if the sample size (n) is large; since then, a single iteration gives satisfactory results. They are not as useful for obtaining the maximum likelihood estimates since each iteration requires solving a complex set of equations. Besides, convergence of this iterative procedure is not guaranteed [Szatrowski (1980)].

A second set of estimating equations is obtained by applying the following results of Magnus (1982).

Lemma 2: Let D be a diagonal T x T matrix with positive diagonal elements and let A and B be square matrices of order k. The kT x kT matrix

$$\Omega = D\,1\,1'D \otimes A + D \otimes B$$

has determinant $|\Omega| = |D|^k\ |C|\ |B|^{T-1}$ and if Ω is non-singular it has inverse

$$\Omega^{-1} = \frac{1}{\alpha}\ 1\,1' \otimes (C^{-1} - B^{-1}) + D^{-1} \otimes B^{-1}$$

where $C = B + \alpha A$ and $\alpha = \operatorname{tr} D$

Lemma 3: For conformal matrices A, B, C, D

$$\operatorname{tr} D\,B'A'C = \operatorname{vec}\,(A')\ (B \otimes C)\ \operatorname{vec}\,(D)$$

Lemma 4: For conformal matrices A, B

$$\operatorname{tr} A'\,B = (\operatorname{vec} A)'\ (\operatorname{vec} B)$$

Applying these results yields

Proposition 2: The maximum likelihood estimates of B, $\ddagger$, and Λ satisfy

$$(\ddagger + \hat{\Lambda}) = \frac{1}{n} \sum_{i=1}^{n} (Y_i - BW_i)\ M(Y_i - BW_i)' \tag{3.13}$$

$$\hat{\ddagger} = [n(T-1)]^{-1} \sum_{i=1}^{n} (Y_i - BW_i)(I-M)(Y_i - BW_i)' \tag{3.14}$$

$$\sum_{i=1}^{n} (Y_i - \hat{B}W_i)W_i' + \sum_{i=1}^{n} T\Lambda\ddagger^{-1}(Y_i - \hat{B}W_i)(I - M)W_i' = 0 \tag{3.15}$$

where $\underset{\sim}{M}$ is the T x T matrix $T^{-1} \underset{\sim}{1} \underset{\sim}{1}'$

The proof of Proposition 2 begins by noting that the Tk x Tk covariance matrix $\underset{\sim}{\Omega}$ can be expressed as

$$|\underset{\sim}{\Omega}| = \underset{\sim}{1}\,\underset{\sim}{1}' \otimes \underset{\sim}{\Lambda} + \underset{\sim}{I}_T \otimes \underset{\sim}{\Sigma}$$

where $\underset{\sim}{1}$ is a T dimensional vector of ones and $\underset{\sim}{I}_T$ is the T x T identity matrix. Applying Lemma 2 yields

$$|\underset{\sim}{\Omega}| = |\underset{\sim}{\Sigma} + T\underset{\sim}{\Lambda}|\ |\underset{\sim}{\Sigma}|^{T-1}$$

and

$$\underset{\sim}{\Omega}^{-1} = T^{-1}\,\underset{\sim}{1}\,\underset{\sim}{1}' \otimes (\underset{\sim}{\Sigma} + T\underset{\sim}{\Lambda})^{-1} + (\underset{\sim}{I} - T^{-1}\,\underset{\sim}{1}\,\underset{\sim}{1}') \otimes \underset{\sim}{\Sigma}^{-1} \tag{3.16}$$

The log of the likelihood function can be expressed

$$\ell = c^* - \frac{n}{2} \log |\underset{\sim}{\Omega}| - \frac{1}{2} \sum_{i=1}^{n} \underset{\sim}{e}_i' \underset{\sim}{\Omega}^{-1} \underset{\sim}{e}_i \tag{3.17}$$

where $c^* = -\frac{1}{2} kT \log (2\pi)$

Using the expression for $\underset{\sim}{\Omega}^{-1}$ displayed in (3.16) and the fact that

$$\underset{\sim}{e}_i = \underset{\sim}{z}_i - (\underset{\sim}{I} \otimes \underset{\sim}{B})\, \underset{\sim}{w}_i$$

yields

$$\underset{\sim}{e}_i' \underset{\sim}{\Omega}^{-1} \underset{\sim}{e}_i = \{[\underset{\sim}{z}_i - (\underset{\sim}{I} \otimes \underset{\sim}{B})\, \underset{\sim}{w}_i]'[\underset{\sim}{M} \otimes (\underset{\sim}{\Sigma} \otimes T\underset{\sim}{\Lambda})^{-1}]$$

$$[\underset{\sim}{z}_i - (\underset{\sim}{I} \otimes \underset{\sim}{B})\, \underset{\sim}{w}_i]\} + \{[\underset{\sim}{z}_i - (\underset{\sim}{I} \otimes \underset{\sim}{B})\, \underset{\sim}{w}_i]'[(\underset{\sim}{I} - \underset{\sim}{M}) \otimes \underset{\sim}{\Sigma}^{-1}]$$

$$[z_i - (\underset{\sim}{I} \otimes \underset{\sim}{B})\, \underset{\sim}{w}_i]\} \tag{3.18}$$

Since $w_i = \text{vec}\,(W_i)$ and $z_i = \text{vec}\,(Y_i)$

$$e_i' \, \Omega^{-1} \, e_i = \{[\text{vec}\,(Y_i) - \text{vec}\,(BW_i)]'[M \otimes (\Sigma + T\Lambda)^{-1}]$$

$$[\text{vec}(Y_i) - \text{vec}\,(BW_i)]\} + \{[\text{vec}(Y_i) - \text{vec}(BW_i)]'$$

$$[(I - M) \otimes \Sigma^{-1}][\text{vec}(Y_i) - \text{vec}\,(BW_i)]\} \qquad (3.19)$$

Plugging (3.19) into (3.17) and applying Lemmas 2 and 3 yields

$$\ell = c^* - \frac{n}{2} \log\,(|\Sigma + T\Lambda||\Sigma|^{T-1}) - \frac{1}{2} \text{tr}[\sum_{i=1}^{n} (Y_i - BW_i)$$

$$M(Y_i - BW_i)'(\Sigma + T\Lambda)^{-1}] - \frac{1}{2} \text{tr}\,[\sum_{i=1}^{n} (Y_i - BX_i)$$

$$(I - M)(Y_i - BW_i)'\Sigma^{-1}]$$

Total differentiation yields

$$d\ell = -\frac{n}{2} \text{tr}\,(\Sigma + T\Lambda)^{-1} \, d(\Sigma + T\Lambda) - \frac{n}{2} (T-1) \, \text{tr}\, \Sigma^{-1} \, d\Sigma$$

$$+ \text{tr} \sum_{i=1}^{n} (\Sigma + T\Lambda)^{-1}(Y_i - BW_i)M \, W_i' \, d\, B' +$$

$$+ \frac{1}{2} \text{tr}[\sum_{i=1}^{n} (\Sigma + T\Lambda)^{-1}(Y_i - BW_i)M(Y_i - BW_i)'(\Sigma + T\Lambda)^{-1}$$

$$(d\Sigma + Td\Lambda)] + \text{tr} \sum_{i=1}^{n} \Sigma^{-1} \, (Y_i - BW_i)(I-M)W_i'dB'$$

$$+ \frac{1}{2} \text{tr} \sum_{i=1}^{n} \Sigma^{-1}(Y_i - BW_i)(I - M)(Y_i - BW_i)'\Sigma^{-1} \, d\,\Sigma$$

which can be written as

$$d\ell = -\frac{1}{2}\, nT \,\mathrm{tr}\, \tilde{\Lambda}\, d\Lambda + \frac{n}{2}\,\mathrm{tr}\, \tilde{\ddagger}\, d\ddagger + \mathrm{tr}\, \tilde{B}\, dB' \tag{3.20}$$

where

$$\tilde{\Lambda} = (\ddagger + T\Lambda)^{-1} + \frac{1}{n}\sum_{i=1}^{n} (\ddagger + T\Lambda)^{-1}(Y_i - BW_i)M(Y_i - BW_i)'(\ddagger + T\Lambda)^{-1}$$

$$\begin{aligned}\tilde{\ddagger} = {} & (\ddagger + T\Lambda)^{-1} - (T-1)\ddagger^{-1} - \frac{1}{n}\sum_{i=1}^{n} [(\ddagger + T\Lambda)^{-1}(Y_i - BW_i)M \\ & (Y_i - BW_i)'(\ddagger + T\Lambda)^{-1}] + \frac{1}{n}\sum_{i=1}^{n} \ddagger^{-1}(Y_i - BW_i)(I - M) \\ & (Y_i - BW_i)'\ddagger^{-1}\end{aligned} \tag{3.21}$$

and

$$\tilde{B} = \sum_{i=1}^{n} (\ddagger + T\Lambda)^{-1}(Y_i - BW_i)\, MW_i' + \sum_{i=1}^{n} \ddagger^{-1}(Y_i - BW_i)(I - M)W_i'$$

Applying Lemma 4 transforms equation (3.20) to

$$d\ell = (nT/2)\ [\mathrm{vec}(d\Lambda)]'\ \mathrm{vec}\tilde{\Lambda} + \frac{n}{2}\ [\mathrm{vec}(d\ddagger)]'\ \mathrm{vec}(\tilde{\ddagger})$$

$$+ [\mathrm{vec}\ (d\, B')]'\ \mathrm{vec}\ \tilde{B}$$

Since $d\Lambda$, $d\ddagger$ and dB are non-zero $d\ell = 0$ implies $\tilde{\Lambda}$, $\tilde{\ddagger}$ and $\tilde{B}$ are 0 which, using the equations displayed in (3.21), yields the three equations (3.13) (3.14) and (3.15).

The two sets of equations we have derived for the maximum likelihood estimates would be of limited value if closed form expressions existed for the maximum likelihood estimators. For completeness, we show that such is not the case - there exists no closed form expression for the maximum likelihood estimates. We apply the following result of Szatrowski (1980)

Lemma 5: For the multivariate linear model with parameter vector $\underset{\sim}{B}$, error covariance matrix $\underset{\sim}{\Omega}$, and design matrix $\underset{\sim}{X}$ there is an explicit solution for the maximum likelihood estimator $\hat{\underset{\sim}{B}}$ of $\underset{\sim}{B}$ if any only if

$$(\underset{\sim}{X}_i' \hat{\underset{\sim}{\Omega}}^{-1} \underset{\sim}{X}_i) \underset{\sim}{X}_i' \hat{\underset{\sim}{\Omega}} = (\underset{\sim}{X}_i' \underset{\sim}{X}_i)^{-1} \underset{\sim}{X}_i$$

We have for our model

Proposition 3: Explicit solutions do not exist, in general, for the maximum likelihood estimates of $\underset{\sim}{B}$ but do exist if $\underset{\sim}{\Lambda} = \underset{\sim}{0}$ in which case the model reduces to the independent error model.

The proof begins by applying Lemma 2 to the covariance matrix $\underset{\sim}{\Omega}$ shown in display (2.4); we have

$$\underset{\sim}{\Omega}^{-1} = \underset{\sim}{M} \otimes (\underset{\sim}{\Sigma} + T\underset{\sim}{\Lambda})^{-1} (\underset{\sim}{I} - \underset{\sim}{M}) \otimes \underset{\sim}{\Sigma}^{-1}$$

where $\underset{\sim}{M} = T^{-1} \underset{\sim}{1} \underset{\sim}{1}'$.

Let $\underset{\sim}{X}_i$ be arranged as the $Tk \times k^2$ matrix $\underset{\sim}{A}_i \otimes \underset{\sim}{I}_k$ where

$$\underset{\sim}{A}_i = \begin{bmatrix} x_{i1}^{(1)} & x_{i1}^{(2)} & \cdots & x_{i1}^{(k)} \\ \vdots & \vdots & & \vdots \\ x_{i,T}^{(1)} & x_{i,T}^{(2)} & \cdots & x_{i,T}^{(k)} \end{bmatrix}$$

where $\underset{\sim}{x}_{it} = (x_{it}^{(1)} \ldots, x_{it}^{(k)})'$

For example, if $k=2$, then the parameter matrix $\underset{\sim}{B} = \begin{matrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{matrix}$

and $\underset{\sim}{X}_i = \underset{\sim}{A}_i \otimes \underset{\sim}{I}$

$$\text{where } \underset{\sim}{A}_i = \begin{bmatrix} x_{i1}^{(1)} & x_{i1}^{(2)} \\ \vdots & \vdots \\ x_{i,T}^{(1)} & x_{i,T}^{(2)} \end{bmatrix}$$

Simple calculation shows

$$\underset{\sim}{X}_i' = \underset{\sim}{A}_i' \otimes \underset{\sim}{I}$$

$$\underset{\sim}{X}_i'\underset{\sim}{\Omega}^{-1} = (\underset{\sim}{A}_i' \otimes \underset{\sim}{I})[\underset{\sim}{M} \otimes (\underset{\sim}{\Sigma} + T\underset{\sim}{\Lambda})^{-1} + (\underset{\sim}{I} - \underset{\sim}{M}) \otimes \underset{\sim}{\Sigma}^{-1}]$$

$$= (\underset{\sim}{A}_i'\underset{\sim}{M}) \otimes (\underset{\sim}{\Sigma} + T\underset{\sim}{\Lambda})^{-1} + \underset{\sim}{A}_i'(\underset{\sim}{I} - \underset{\sim}{M}) \otimes \underset{\sim}{\Sigma}^{-1}$$

consequently

$$\underset{\sim}{X}_i'\underset{\sim}{\Omega}^{-1}\underset{\sim}{X}_i = (\underset{\sim}{A}_i'\underset{\sim}{M}\underset{\sim}{A}_i) \otimes (\underset{\sim}{\Sigma} + T\underset{\sim}{\Lambda})^{-1} + \underset{\sim}{A}_i' (\underset{\sim}{I} - \underset{\sim}{M})\underset{\sim}{A}_i \otimes \underset{\sim}{\Sigma}^{-1} \qquad (3.22)$$

$$(\underset{\sim}{X}_i'\underset{\sim}{X}_i)^{-1} \underset{\sim}{X}_i = (\underset{\sim}{A}_i'\underset{\sim}{A}_i \otimes \underset{\sim}{I})^{-1} (\underset{\sim}{A}_i' \otimes \underset{\sim}{I})$$

$$= (\underset{\sim}{A}_i'\underset{\sim}{A}_i)^{-1} \underset{\sim}{A}_i \otimes \underset{\sim}{I}$$

and thus

$$(X_i'\Omega^{-1}X_i)(X_i'X_i)^{-1}X_i = A_i'MA_i(A_i'A_i)^{-1}A_i' \otimes (\Sigma + T\Lambda)^{-1}$$

$$+ A_i'(I - M)A_i(A_i'A_i)^{-1}A_i' \otimes \Sigma^{-1} \qquad (3.23)$$

Suppose the expressions in (3.22) and (3.23) are equivalent. Since $(\Sigma + T\Lambda)$ and Σ are positive definite either a) $\Lambda = 0$ or b) $A_i(A_i'A_i)^{-1}A_i' = I$. If $\Lambda = 0$ the model becomes the independent errors model studied by Mayer (1985) and Anderson (1978) which does allow an explicit solution for $\hat{B}$. If for every i, $A_i(A_i'A_i)^{-1}A_i' = I$ then $X_i\hat{b}_i = X_i(X_i'X_i)^{-1}X_i'z_i = z_i$ and the model is degenerate. Since neither condition holds in general, applying Lemma 5 yields that there is no explicit maximum likelihood estimator.

Having shown that no closed form expression exists, in general, for the maximum likelihood estimators, we display a modified, somewhat simplified, model for which the maximum likelihood estimators yield closed-form expression. This result extends a result of Anderson and Hsiao (1981) to the multivariate panel model.

<u>Proposition 4</u>: Explicit maximum likelihood estimators of B do exist for the model displayed in (2.1) and (2.2) provided there exists an additional wave 0 for which z_{i0} has structure

$$z_{i0} = B\, w_{i0} + \alpha_i \qquad i = 1,\dots,n$$

This model assumes that the only error for the initial observation is the random individual effect α_i.

For this model equations (2.1) and (2.2) can be combined into

$$\underset{\sim}{z}_{it} = \underset{\sim}{B}\underset{\sim}{w}_{it} + (\underset{\sim}{z}_{i0} - \underset{\sim}{B}\underset{\sim}{w}_{i0}) + \underset{\sim}{u}_{it}$$

The likelihood function can be expressed as

$$L(\underset{\sim}{B}, \underset{\sim}{\Lambda}, \underset{\sim}{\Sigma}) = (2\pi)^{-nk(T+1)/2} \quad |\underset{\sim}{\Lambda}|^{-n/2} \; |\underset{\sim}{\Sigma}|^{-nT/2}$$

$$\exp\left(-\frac{1}{2}\right) \{\sum_{t=1}^{T} \mathrm{tr}\, \underset{\sim}{\Sigma}^{-1} \sum_{i=1}^{n} [\underset{\sim}{z}_{it} - \underset{\sim}{z}_{i0} - \underset{\sim}{B}\,(\underset{\sim}{w}_{it} - \underset{\sim}{w}_{i0})]$$

$$[\underset{\sim}{z}_{it} - \underset{\sim}{z}_{i0} - \underset{\sim}{B}(\underset{\sim}{w}_{it} - \underset{\sim}{w}_{i0})]' + \mathrm{tr}\, \underset{\sim}{\Lambda}^{-1} \sum_{i=1}^{n} \underset{\sim}{z}_{i0}\underset{\sim}{z}_{i0}'\}$$

If we let $\ell = \frac{2}{n} \log L$ then

$$\ell = c_1 - \log |\underset{\sim}{\Lambda}| - T \log |\underset{\sim}{\Sigma}| - \frac{1}{n} \{ \sum_{t=1}^{T} \mathrm{tr}\, \underset{\sim}{\Sigma}^{-1} \sum_{i=1}^{n}$$

$$[\underset{\sim}{z}_{it} - \underset{\sim}{z}_{i0} - \underset{\sim}{B}(\underset{\sim}{w}_{it} - \underset{\sim}{w}_{i0})][\underset{\sim}{z}_{it} - \underset{\sim}{z}_{i0} - \underset{\sim}{B}(\underset{\sim}{w}_{it} - \underset{\sim}{w}_{i0})]' +$$

$$\mathrm{tr}\, \underset{\sim}{\Lambda}^{-1} \sum_{i=1}^{n} \underset{\sim}{z}_{i0}\underset{\sim}{z}_{i0}' \tag{3.24}$$

where $c_1 = - k\,(T + 1) \log\,(2\pi)$

To maximize the expression in (3.24) we take derivatives and set them equal to zero yielding

$$\frac{\partial \ell}{\partial \underset{\sim}{B}} = \frac{2}{n} \sum_{t=1}^{T} \underset{\sim}{\Sigma}^{-1} \sum_{i=1}^{n} [\underset{\sim}{z}_{it} (\underset{\sim}{w}_{it} - \underset{\sim}{w}_{i0})' - \underset{\sim}{z}_{i0} (\underset{\sim}{w}_{it} - \underset{\sim}{w}_{i0})'$$

$$- \underset{\sim}{B}(\underset{\sim}{w}_{it} - \underset{\sim}{w}_{i0})(\underset{\sim}{w}_{it} - \underset{\sim}{w}_{i0})'] = \underset{\sim}{0}$$

which yields the maximum likelihood estimator

$$\hat{\underset{\sim}{B}} = \{(nT)^{-1} \sum_{t=1}^{T} \sum_{i=1}^{n} \underset{\sim}{z}_{it} (\underset{\sim}{w}_{it} - \underset{\sim}{w}_{i0})' - (\underset{\sim}{w}_{it} - \underset{\sim}{w}_{i0}) \underset{\sim}{z}_{i0}'\}$$

$$\{(nT)^{-1} \sum_{t=1}^{T} \sum_{i=1}^{n} (\underset{\sim}{w}_{it} - \underset{\sim}{w}_{i0})(\underset{\sim}{w}_{it} - \underset{\sim}{w}_{i0})'\}^{-1}$$

For the maximum likelihood estimator of the covariance matrices we have

$$\frac{\partial \ell}{\partial \underset{\sim}{\Lambda}} = - \underset{\sim}{\Lambda}^{-1} - n^{-1} [- \underset{\sim}{\Lambda}^{-1} \underset{\sim}{\Lambda}'^{-1} \sum_{i=1}^{n} \underset{\sim}{z}_{i0} \underset{\sim}{z}_{i0}'] = 0$$

which yields the maximum likelihood estimator

$$\hat{\underset{\sim}{\Lambda}} = \frac{1}{n} \sum_{i=1}^{n} \underset{\sim}{z}_{i0} \underset{\sim}{z}_{i0}'$$

and, finally we have

$$\frac{\partial \ell}{\partial \underset{\sim}{\Sigma}} = - T \underset{\sim}{\Sigma}^{-1} - n^{-1} \{ \sum_{t=1}^{T} (- \underset{\sim}{\Sigma}^{-1} \underset{\sim}{\Sigma}^{-1}) \sum_{i=1}^{n} [\underset{\sim}{z}_{it} - \underset{\sim}{z}_{i0} - \underset{\sim}{B}(\underset{\sim}{w}_{it} - \underset{\sim}{w}_{i0})]$$

$$[\underset{\sim}{z}_{it} - \underset{\sim}{z}_{i0} - \underset{\sim}{B}(\underset{\sim}{w}_{it} - \underset{\sim}{w}_{i0})]'\} = \underset{\sim}{0}$$

which yields the maximum likelihood estimator

$$\hat{\underset{\sim}{\Sigma}} = (nT)^{-1} \{ \sum_{t=1}^{T} \sum_{i=1}^{n} (\underset{\sim}{z}_{it} - \underset{\sim}{z}_{i0} - \underset{\sim}{B}(\underset{\sim}{w}_{it} - \underset{\sim}{w}_{i0})]$$

$$[\underset{\sim}{z}_{it} - \underset{\sim}{z}_{i0} - \underset{\sim}{B}(\underset{\sim}{w}_{it} - \underset{\sim}{w}_{i0}]'\}$$

which completes the proof.

Returning to the more general model, we consider a strategy for computing the maximum likelihood estimates. It begins by computing the maximum likelihood estimate of $\underset{\sim}{B}$ for the model without constraining the entries of $\underset{\sim}{\Omega}$ to have exactly the structure displayed in (2.4). Instead, $\underset{\sim}{\Omega}$ is replaced by $\underset{\sim}{\Omega}^*$ where

$$\underset{\sim}{\Omega}^* = \begin{bmatrix} \underset{\sim}{\Psi}^* & & & \underset{\sim}{\Lambda}^* \\ & \cdot & & \\ & & \cdot & \\ \underset{\sim}{\Lambda}^* & & \cdot & \underset{\sim}{\Psi}^* \end{bmatrix}$$

has identical k x k diagonal blocks, $\underset{\sim}{\Psi}^*$, and identical off-diagonal blocks $\underset{\sim}{\Lambda}^*$. In the original notation $\underset{\sim}{\Psi}^* = \underset{\sim}{\Sigma} + \underset{\sim}{\Lambda}$ but the constraint that $\underset{\sim}{\Psi}^* - \underset{\sim}{\Lambda}^*$ is positive definite is not enforced. The advantage of this modified model is that the calculations are performed by a multivariate regression program or a simple covariance structure routine such as the LISREL routine.

Once the estimates $\hat{\underset{\sim}{B}}^*$, $\hat{\underset{\sim}{\Psi}}^*$ and $\hat{\underset{\sim}{\Lambda}}^*$ are obtained the eigenvalues of $\hat{\underset{\sim}{\Psi}}^* - \hat{\underset{\sim}{\Lambda}}^*$ are examined. If they are all positive then $\hat{\underset{\sim}{\Sigma}} = \hat{\underset{\sim}{\Psi}}^* - \hat{\underset{\sim}{\Lambda}}^*$,

$\hat{\mathbf{B}} = \hat{\mathbf{B}}$ and $\hat{\mathbf{\Lambda}} = \hat{\mathbf{\Lambda}}$ are the maximum likelihood estimates for the panel model with random individual effects.

If the eigenvalues of $\mathbf{\Psi}^* - \mathbf{\Lambda}^*$ are not all positive we apply the following method, due to Klotz and Putter (1969) and used by Anderson (1984, 1985), obtain the maximum likelihood estimates $\hat{\mathbf{B}}$, $\hat{\mathbf{\Sigma}}$ and $\hat{\mathbf{\Lambda}}$.

Let $\mathbf{Z}$ be a nonsingular matrix and $D = \text{diag}\ (d_1,\ldots,d_k)$ a diagonal matrix satisfying

$$\hat{\mathbf{\psi}}^* = \mathbf{ZDZ}' \text{ and } \hat{\mathbf{\Lambda}}^* = \mathbf{ZZ}'$$

Let k^* be the number of d_i satisfying $d_i > 1$ and let D^* be the k^* diagonal matrix of these d_i. Let $\mathbf{Z}^*$ be the matrix of k^* rows of $\mathbf{Z}$, the rows corresponding to those elements. The maximum likelihood estimate of $\mathbf{\Sigma}$ is

$$\hat{\mathbf{\Sigma}} = (\mathbf{Z}^*\mathbf{D}^*\mathbf{Z}^{*\prime} - \mathbf{Z}^*\mathbf{Z}^{*\prime})$$

and the maximum likelihood estimator of $\mathbf{\Lambda}$ is

$$\hat{\mathbf{\Lambda}} = \hat{\mathbf{\Psi}}^* - \hat{\mathbf{\Sigma}}$$

The maximum likelihood estimator of $\mathbf{B}$ is obtained by constraining $\mathbf{\Lambda} = \hat{\mathbf{\Lambda}}$ and $\mathbf{\Sigma} = \hat{\mathbf{\Sigma}}$ and applying LISREL or any generalized least squares multivariate regression procedure.

In anticipation of our numerical example two other estimators of $\underset{\sim}{B}$ deserve mention. First, the model can be transformed to differences thus eliminating, or masking, the individual effects. The "pseudo-maximum likelihood" estimator of $\underset{\sim}{B}$ is the maximum likelihood estimator for the transformed model

$$\underset{\sim}{z}_{it} - \underset{\sim}{z}_{i,t-1} = \underset{\sim}{B}(\underset{\sim}{w}_{it} - \underset{\sim}{w}_{i,t-1}) + (\underset{\sim}{e}_{it} - \underset{\sim}{e}_{i,t-1}) \qquad t = 1,\ldots,T$$

This estimator is consistent in n and is asymptotically efficient. In practice this method seems to yield good approximations for the regression coefficients, poor approximations for the estimates of the $\underset{\sim}{\Sigma}$ matrix and no approximation for the $\underset{\sim}{\Lambda}$ matrix. We do not suggest this method be used. Similarly there are several instrumental variable methods have been introduced for the univariate version of this model with lagged endogenous predictors [Anderson and Hsiao (1981), Amemiya and MaCurdy (1986), Breush, Mizon, and Schmidt (1986)]. These estimators can be extended to the individual effects model but in light of the absence of lagged predictors and the ease of the Anderson one-step estimate we do not pursue these alternatives further.

Second the individual random component of the error could be ignored and estimators appropriate for the independent error model be used. These "multivariate regression estimators" are obtained from the pooled multivariate regression of $\underset{\sim}{z}_{it}$ on $\underset{\sim}{w}_{it}$. These estimators ignore the individual effect but are consistent in the number of observations. Moreover, if the correlation between the errors of the dependent variables is also ignored then "multiple regression estimators"

of $\underset{\sim}{B}$ and $\underset{\sim}{\ddagger}$ can be obtained by isolating and regressing each endogenous variable on all exogenous variables.

4. Application to a Study of Perception of Care given by HMO's

To illustrate our methods of estimation we consider a panel study conducted by a consulting firm of the relationship between patient's attitudes toward Health Maintenance Organizations (HMO) and their perceptions of the quality of care given by such organizations. These data are proprietary and have been used to illustrate methods of testing the strength effects in multivariate cross-lagged panel models [Mayer (1985), Mayer and Carroll (1986)]. Here the analysis is used to illustrate our methods of estimating the effects of the attitudes on the perceptions over waves. We formulate the appropriate panel model and then estimate and interpret the regression coefficients and the variance and covariance terms.

The data were obtained from interviews of 100 randomly selected patients upon entering and upon leaving a single HMO after having requested and seen a primary care physician for the first time since enrolling in the HMO. The same 100 patients are interviewed again, before and after their next three visits to the HMO. The 100 patients were picked to have enrolled in the HMO at about the same time and to have approximately the same length of time between visits - about 6 months. The visits are for themselves, not for family members (e.g., pediatric care or geriatrics care), and are for primary care, as opposed to specialized care. None of the patients are pregnant, seriously ill,

chronically ill, or mentally ill and none have suffered serious injuries.

The four variables of interest are each a compilation of several measurements. The raw data were not provided by the firm that collected the data. We note that the use of a single HMO makes the analysis highly tentative in that it restricts the external validity of the findings. However before undertaking a more complete study the firm wanted to know if such effects existed in these data.

The first predictor variable, x_1, measures the degree to which the patient has a positive attitude toward the concept of an HMO as a provider of primary medical care; it is measured at each visit when the patient arrives for his or her appointment. Issues of secondary care are not addressed. The second predictor variable, x_2, measures the degree to which the patient has a positive attitude about the specific HMO in which he or she in enrolled; it is measured at each visit after the patient has seen the physician, and completed any treatments.

The first dependent variable, z_1, samples the patient's perception of the quality of care given by HMO's in general and the second dependent variables, z_2, samples the patient's perception of the quality of care given by the particular HMO in which he or she is enrolled. Both are assessed after the patient has seen the physician and completed any treatments.

All of the variables were scaled to have mean 10 and standard deviation of about 1. This scaling protects the confidentiality of the data and does not affect our modeling since the measurements have no natural scale.

The model adopted is the individual effect model displayed in (2.1) and (2.2). The parameters are estimated in each of four ways:

A. The maximum likelihood method

B. The multivariate regression method (ignores individual effect)

C. The multiple regression method (ignores individual effect)

D. The one-step iteration method of Anderson

E. The pseudo-maximum likelihood method

The correlation statistics for the raw data are presented in Table 1 and the results of estimating the model by maximum likelihood are summarized in Table 2.

We begin interpreting the model estimated by maximum likelihood by noting that the chi-square goodness of fit measure for the model is 461.22 with 90 degrees of freedom, indicating that the model fits the data rather poorly. In light of this lack of fit we treat all further analysis as quite preliminary and somewhat speculative.
The largest effect over time - the coefficient is .162 - is the effect of the patient's attitude toward the concept of an HMO as a provider of health-care on the perception of the quality of care given by the specific HMO in which he or she is enrolled. The marginal t statistic -

computed from the estimate of the standard error of the coefficient obtained from the information matrix - is 4.731 indicating a that the coefficient is quite different from zero. As a guide to interpretation we tend to use the .05 significance level and two-tailed rejection regions but because of the limited sample size we do no formal hypothesis testing. Patients that are positive toward HMO's as health-care providers tend to perceive that their HMO gives high quality care. On the other hand, the same attitude does not appear to have as strong of effect on the patient's perception of the quality of care given by HMO's in general - the regression coefficient is -.013 with a t statistic of -.327.

Similarly, the patient's attitude toward the particular HMO in which he or she is enrolled appears to have a very small effect, if any effect at all, on either the perception of the quality of care given by HMO's in general or on the perception of the quality of care given by the particular HMO in which the patient is enrolled - regression coefficients of .056 and .053 with t statistics of 1.38 and 1.54.

These results, if found to be obtain across a wide range of samples, may be important to the HMO industry as they desire to understand correlates of patient's perceptions of the quality of care given by HMO's in general and by particular HMO's, in order, in part, to enhance the attractiveness of the HMO to workers asked to choose between an HMO and alternative programs such as third-party insurance that covers part of the cost of fee-for-service medicine. If workers make their choice, in part, based on their perception of the quality of care being given by

the particular HMO being offered and if that perception can be improved by improving the general attitude toward HMO's then it may be wise to invest in such improvement. On the other hand, it may be less wise to invest in improving the attitude toward the particular HMO being offered since that attitude, at least in these data, is not highly related to the perception of quality of the care being offered.

Similarly, our analysis suggests that investing in improving either of the attitude measures may have little or no effect on the perception of the quality of care given by HMO's in general. It may not be fruitful to attempt to improve this perception by investing in improving the attitude toward HMO's in general or toward the specific HMO. Although more studies would be needed before any firm conclusions could be reached our model does suggest hypotheses worthy of further study.

Considering the maximum likelihood estimates of the two covariance matrices, $\underset{\sim}{\Sigma}$ and $\underset{\sim}{\Lambda}$, we note that the entries in $\underset{\sim}{\hat{\Lambda}}$ are considerably larger than their estimated standard errors and in two of the three cases are larger than the corresponding entries in $\underset{\sim}{\hat{\Sigma}}$, their size seems to indicate that the individual effects may be a larger source of variation in the data than are the pure errors. We find this pattern frequently in fitting this type of model. If the individual effect is present in the data but ignored in the model its variance tends to be added to the estimate of the covariance matrix of the pure error and may lead to considerable over-estimation of the pure error in the model as we will se below.

Turning to the comparison of the maximum likelihood estimators with the other estimators - whose values are summarized in Table 3 - we note considerable similarities with a few noteworthy exceptions. The first estimates in Table 3 are those obtained by ignoring the individual effects and fitting the independent errors model by multivariate regression methods. The regression coefficients are quite similar to those obtained by maximum likelihood except the effect of x_1 on z_1 changes sign, going from -.013 to .110 with the t-statistic going from -.327 to 2.11 indicating a significant positive effect of attitude toward HMO's in general on the perception of the quality of care given by HMO's in general.

Most interestingly, the estimate of the covariance estimate obtained from fitting the independent error model approximates the sum of the estimates of the two covariance matrices, $\underset{\sim}{\ddagger}$ and $\underset{\sim}{\Lambda}$ in the model which allows an individual effect. For example, the estimate of the pure error of the attitude toward HMO's in general is .441 compared to .139 for the individual effect model.

The chi square goodness of fit statistic for the independent error model fit by multivariate regression methods is 786 with 93 degrees of freedom (compared to 461 with 90 degrees of freedom for the model with individual effects) indicating that removing the individual effect reduces the goodness of fit sharply. Even more critically, the t statistics obtained from the multivariate regression calculations are different than the t statistics for the model with individual effects. Most notabley, as indicated in Table 3, the effect of the attitude

toward HMO's in general appears to have a fairly strong positive effect on the perception of quality of care given by HMO's in general. This effect appears to be inflated by the exclusion of the individual effect.

The estimators obtained for the independent errors model by applying multiple regression methods to each equation separately are identical to those obtained from the multivariate regression methods except the off-diagonal entry in $\underset{\sim}{\Psi}$, the covariance matrix of the errors is estimated, apriori, to be zero. The goodness of fit statistic for this model is 818.74 with 94 df, a further large reduction in the quality of fit.

The Anderson one-step estimates are good approximations to the maximum likelihood estimates and this closeness has held across a number of applications. In our experience these estimates are preferable to those obtained from applying independent errors model.

The pseudo maximum likelihood estimates obtained from the equations for differences are fairly similar to the original maximum likelihood estimates but are quite inferior to the Anderson one-step estimates. The pseudo maximum likelihood method provides no estimate of the matrix $\underset{\sim}{\Lambda}$.

As for the issue of computation, the Anderson one-step estimators are very similar to those obtained by maximum likelihood and involve relatively simple calculations. We suggest anyone able to do a generalized least-squares regression may want to consider using his

simple approximation rather than calculating the more complicated maximum likelihood estimates. The similarity between the estimates given by his method and the maximum likelihood method has obtained over a variety of panel examples we have considered.

5. Discussion

We have considered the problem of maximum likelihood estimation for a continuous variable panel model characterized by multiple responses, multiple predictors and an individual specific random component in the error structure. We show several representations for these estimators, show that the can not be written, in general, in closed form, and then show that they can be written in closed form for a special case of the model.

Our results are intended to help those empirical social scientists trying to model panel studies but are skeptical of assuming the errors for a single individual are independent over time. They are also intended to help those empirical social scientists that would use a covariance structure approach to obtain maximum likelihood estimates of the parameters but are curious to know more about the estimates they have obtained.

We apply our results to a set of data on the relationship between the perceptions and attitudes of patients enrolled in a health maintenance organization and show that the attitude toward HMO's in general may be correlated with the perception of quality of care given by a specific

HMO but that little correlation appears to exist between attitude toward HMO's in general and perception of the general quality of care given by HMO's. We show that the individual effect may account for more random variation than does the pure error and that ignoring this effect, by fitting the independent errors model by multivariate or multiple regression, may give somewhat misleading results.

Our results need to be extended in several directions. First similar study of maximum likelihood estimation is needed for the model with lagged endogenous predictors. Second, for the model with such lagged predictors the problem of testing for the presence of cross-effects, a central issue in many applications of continuous variable panel methods, needs to be addressed. The model with latent variables which are imperfectly reflected by empirical measures should be explored and finally the properties of the alternative estimators dealing with samples of moderate size should be considered.

Table 1: The raw correlations for the HMO data

The simple correlations:

	x_1	x_2	z_1
x_2	.298		
z_1	.032	.030	
z_2	.368	.189	.093

The multiple correlations:

$[R(z_1; x_1, x_2)] = .001$

$[R(z_2; x_1, x_2)] = .143$

Table 2. Maximum likelihood estimates for the HMO data

A. Maximum likelihood estimators

$$\underset{\sim}{B}: \begin{matrix} -.013 & .056 \\ .162 & .053 \end{matrix} \qquad \underset{\sim}{\Sigma}: \begin{matrix} .139 & .015 \\ .015 & .110 \end{matrix} \qquad \underset{\sim}{\Lambda}: \begin{matrix} 309 & .076 \\ .076 & .112 \end{matrix}$$

B. Estimated standard errors

$$\underset{\sim}{B}: \begin{matrix} .041 & .041 \\ .034 & .035 \end{matrix} \qquad \underset{\sim}{\Sigma} + \underset{\sim}{\Lambda}: \begin{matrix} .052 & .025 \\ .025 & .022 \end{matrix} \qquad \underset{\sim}{\Lambda}: \begin{matrix} .051 & .024 \\ .024 & .021 \end{matrix}$$

C. Asymtotic (marginal) t statistics

$$\underset{\sim}{B}: \begin{matrix} -.327 & 1.38 \\ 4.731 & 1.54 \end{matrix}$$

Chi-square goodness of fit statistic = 461 df = 90

Table 3: Alternative estimators of the coefficients for the HMO data

A. Independent errors model: Multivariate regression estimators

$$\underset{\sim}{B}: \begin{matrix} .110 & .057 \\ .245 & .066 \end{matrix} \qquad \underset{\sim}{\Sigma}: \begin{matrix} .441 & .086 \\ .086 & .218 \end{matrix} \qquad \underset{\sim}{\Lambda}: \text{No estimate}$$

$$\text{t statistics for regression coefficients} = \begin{matrix} 2.11 & 1.02 \\ 6.69 & 1.67 \end{matrix}$$

Chi-square goodness of fit = 786 df = 93

B. Independent errors model: Multiple regression estimators

Same as A. except the estimated covariance $\hat{\sigma}_{12} = 0$ and

Chi-square goodness of fit = 819 df = 94

C. Anderson one-step estimators

$$\underset{\sim}{B}: \begin{matrix} -.009 & .056 \\ .166 & .053 \end{matrix} \qquad \underset{\sim}{\Sigma}: \begin{matrix} .151 & .020 \\ .020 & .113 \end{matrix} \qquad \underset{\sim}{\Lambda}: \begin{matrix} .292 & .065 \\ .065 & .103 \end{matrix}$$

D. Psuedo maximum-likelihood estimators

$$\underset{\sim}{B}: \begin{matrix} -.012 & .040 \\ .058 & -.015 \end{matrix} \qquad \underset{\sim}{\Sigma}: \begin{matrix} .073 & .003 \\ .003 & .087 \end{matrix} \qquad \underset{\sim}{\Lambda}: \text{No estimate}$$

REFERENCES

Amemiya, T. and McCurdy, T. (1986). Instrumental-Variable Estimation of An Error-Components Model, Econometrica, 54, 869-880.

Anderson, T. W. (1969). Statistical Inference for Covariance Matrices with Linear Structure, in Proceedings of the Second International Symposium on Multivariate Analysis, (P. R. Krishnaiah, ed.), New York: Academic Press, 55-66.

Anderson, T. W. (1970). Estimation of Covariance Matrices Which Are Linear Combinations or Whose Inverses Are Linear Combinations of Given Catrices, in Essays in Probability and Statistics, (R. C. Bose, et. al, eds.), Chapel Hill: University of North Carolina Press, 1-24.

Anderson, T. W. (1973). Asymptotic Efficient Estimation of Covariance Matrices with Linear Structure, The Annals of Statistics, 1, 135-141.

Anderson, T. W. (1978). Repeated Measurements on Autoregressive processes. Journal of the American Statiscal Association, 73, 371-378.

Anderson, T. W. (1984). Estimating Linear Statistical Relationships, The 1982 Wald Memorial Lectures, The Annals of Statistics, 12 (1), 1-45.

Anderson, T. W. (1985). Components of Variance in MANOVA, Multivariate Analysis - VI (P. R. Krishnaiah, ed.), Amsterdam: Elsevier Science Publishers, 1-8.

Anderson, T. W. and Hsiao, C. (1981). Estimation of Dynamic Models with Error Components, Journal of the American Statistical Association, 76, 598-606.

Anderson, T. W. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models Using Panel Data, Journal of Econometrics, 18, 47-82.

Balestra, P. and Nerlove, M. (1966). Pooling Cross Section and Time Series Data in the Estimation of a Dynamic Model: The Demand for Natural Gas, Econometrica, 34, 585-612.

Breush, T. S. Mizon, G. E. and Schmidt, P. (1986). Some Results on Panel Data, Unpublished paper.

Chamberlain, G. (1984). Panel Data, in Handbook of Econometrics, Vol. II, (Z. Griliches and M. Intriligator, eds.), Amsterdam: Elsevier Science Publishers, 1247-1318.

Dielman, T. E. (1983). Pooled Cross-Sectional and Time Series Data: A Survey of Current Statistical Methodology, The American Statistican, 37, 111-122.

Hausman, J. A. and Taylor, W. E. (1981). Panel Data and Unobservable Individual Effects, Econometrica, 49, 1377-1398.

Hsiao, C. (1986). Analysis of Panel Data, Cambridge: Cambridge University Press.

Hsiao, C. (1985). Benefits and Limitations of Panel Data, Econometric Reviews, 4, 121-174.

Joreskog, K. G. and Goldberger, A. G. (1975). Estimation of a Model with Multiple Indicators and Multiple Causes of a Single Latent Variable, Journal of the American Statistical Association, 70, 631-639.

Joreskog, K. G. and Sorbom, D. (1984). LISREL VI Analysis of Linear Structural Relationships by the Method of Maximum Likelihood, Mooresville: Scientific Software.

Klotz, J. and Putter, J. (1969). Maximum Likelihood Estimation of Multivariate Covariance Components for The Balanced One-Way Layout, The Annals of Mathematical Statistics 40 (3), 1100-1105.

Magnus, J. R. (1982). Multivariate Error Components Analysis of Linear and Nonlinear Regression Models by Maximum Likelihood, Journal of Econometrics, 19, 239-285.

Mayer, L. S. (1985). Statistical Inferences in Cross-Lagged Panel Studies, Technical Report, Department of Statistics, Standford University.

Mayer, L. S. (1986). Statistical Inferences for Cross-Lagged Panel Models without the Assumption of Normal Errors, Social Science Research, 15, 28-42.

Mayer, L. S. (1986). On Cross-Lagged Panel Models with Serially Correlated Errors, Journal of Business and Economic Statistics, 4(3), 347-357.

Mayer, L. S. and Carroll S. S. (1986). Detecting Serial Correlation in the Error Structure of a Cross-Lagged Panel Model, Communications in Statistics - Theory and Methods, 15 (2), 345-366.

Neudecker, H. (1969). Some Theorems on Matrix Differentiation with Special Reference to Kronecker Matrix Products, Journal of the American Statistical Association, 64, 953-963.

Szatrowski, T. H. (1980). Necessary and Sufficient Conditions for Solutions in the Multivariate Normal Estimation Problem for Patterned Means and Covariances, The Annals of Statistics, 8 (4), 802-810.

Technical Reports

U.S. Army Research Office - Contracts DAAG29-82-K-0156 and DAAG29-85-K-0239

1. "Maximum Likelihood Estimators and Likelihood Ratio Criteria for Multivariate Elliptically Contoured Distributions," T. W. Anderson and Kai–Tai Fang, September 1982.
2. "A Review and Some Extensions of Takemura's Generalizations of Cochran's Theorem," George P.H. Styan, September 1982.
3. "Some Further Applications of Finite Difference Operators," Kai–Tai Fang, September 1982.
4. "Rank Additivity and Matrix Polynomials," George P.H. Styan and Akimichi Takemura, September 1982.
5. "The Problem of Selecting a Given Number of Representative Points in a Normal Population and a Generalized Mills' Ratio," Kai–Tai Fang and Shu–Dong He, October 1982.
6. "Tensor Analysis of ANOVA Decomposition," Akimichi Takemura, November 1982.
7. "A Statistical Approach to Zonal Polynomials," Akimichi Takemura, January 1983.
8. "Orthogonally Invariant Minimax Estimator of the Covariance Matrix of a Multivariate Normal Population," Akimichi Takemura, April 1983.
9. "An Orthogonally Invariant Minimax Estimator of the Covariance Matrix of a Multivariate Normal Population," Akimichi Takemura, April 1983.
10. "Relationships Among Classes of Spherical Matrix Distributions," Kai–Tai Fang and Han–Feng Chen, April 1984.
11. "A Generalization of Autocorrelation and Partial Autocorrelation Functions Useful for Identification of ARMA(p,q) Processes," Akimichi Takemura, May 1984.
12. "Methods and Applications of Time Series Analysis Part II: Linear Stochastic Models," T. W. Anderson and N. D. Singpurwalla, October 1984.

13. "Why Do Noninvertible Estimated Moving Averages Occur?" T. W. Anderson and Akimichi Takemura, November 1984.

14. "Invariant Tests and Likelihood Ratio Tests for Multivariate Elliptically Contoured Distributions," Huang Hsu, May 1985.

15. "Statistical Inferences in Cross–lagged Panel Studies," Lawrence S. Mayer, November 1985.

16. "Notes on the Extended Class of Stein Estimators," Suk–ki Hahn, July 1986.

17. "The Stationary Autoregressive Model," T. W. Anderson, July 1986.

18. "Bayesian Analyses of Nonhomogeneous Autoregressive Processes," T. W. Anderson, Nozer D. Singpurwalla, and Refik Soyer, September 1986.

19. "Estimation of a Multivariate Continuous Variable Panel Model," Lawrence S. Mayer and Kun Mao Chen, June 1987.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 19	2. GOVT ACCESSION NO. N/A	3. RECIPIENT'S CATALOG NUMBER N/A
4. TITLE (and Subtitle) ESTIMATION OF A MULTIVARIATE CONTINUOUS VARIABLE PANEL MODEL		5. TYPE OF REPORT & PERIOD COVERED Technical
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Lawrence S. Mayer and Kun Mao Chen		8. CONTRACT OR GRANT NUMBER(s) DAAG29-85-K-0239
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Statistics - Sequoia Hall Stanford University Stanford, California 94305-4065		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS P-19065-M
11. CONTROLLING OFFICE NAME AND ADDRESS U. S. Army Research Office P. O. Box 12211 Research Triangle Park, NC 27709		12. REPORT DATE June 1987
		13. NUMBER OF PAGES 37 pp.
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

N/A

18. SUPPLEMENTARY NOTES

The view, opinions, and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other documentation.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Panel Analysis; Components of Variance; Linear Models: Maximum-likelihood.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

(See reverse side for Abstract.)

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

20. Abstract.

Continuous-variable panel models are widely used in social and business research to assess the effects of exogenous (and lagged endogenous variables) on endogenous variables over time given a large number of independent replications on variables measured at each of a small number of times. We consider the problem of estimating the parameters for a model in which the error includes a random individual component. The univariate version of this "components of variance" model has been widely studied in econometrics. Results in the mathematical statistics literature are used to characterize and study the maximum-likelihood estimators. The results are applied to a panel study of patients' attitudes and perceptions toward health maintenance organizations.